DOCUMENT RESUME

ED 435 683 TM 030 314

AUTHOR Davila, Norma

TITLE Charting the Future of Assessment in Systemic Educational Reform: Teacher and School Principal Involvement in Evaluation and Assessment Use.

PUB DATE 1999-11-00

NOTE 14p.; Paper presented at the Annual Meeting of the American Evaluation Association (14th, Orlando, FL, November 3-6, 1999). Adaptation of a paper presented at the Annual NISE Forum (4th, 1999).

PUB TYPE Reports - Descriptive (141) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS *Academic Achievement; *Educational Change; Educational Research; Elementary Secondary Education; Evaluation Methods; Evaluation Utilization; *Outcomes of Education; Participation; Systems Development; *Teachers

IDENTIFIERS *Puerto Rico; *Reform Efforts; Statewide Systemic Initiative

ABSTRACT

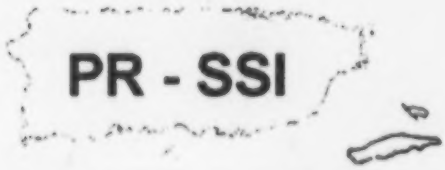
This paper presents the evolution of and the lessons learned from a research approach to assessment of student outcomes, specifically of student academic achievement, being used by the Puerto Rico Statewide Systemic Initiative (PR-SSI), which is one of the statewide systemic initiatives for science and mathematics sponsored by the National Science Foundation. A participatory-research approach was selected for the evaluation and assessment of the PR-SSI in general and for the assessment of student academic achievement in particular. Triangulation of results has been a major element of the design from the beginning of the reform. Two different versions of the PR-SSI's model were developed. The evolution of the first and second versions show that considerable organizational learning took place within the PR-SSI. The final model, which incorporates information from the Third International Mathematics and Science Study and the engagement of teachers and administrators, challenges traditional notions of participant involvement in evaluation because it requires direct contributions to use results to improve teaching and learning and the teamwork of evaluators and content specialists to lead the self-assessment process. (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 435 683

PRDE

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Norma Dávila
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)



RCSE COMM

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Charting the Future of Assessment in Systemic Educational Reform: Teacher and School Principal Involvement in Evaluation and Assessment Use

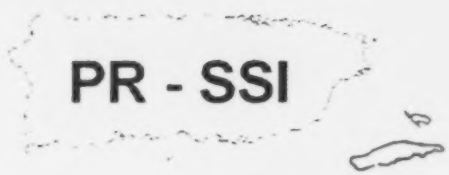
Norma Dávila
University of Puerto Rico

Paper Presented at the 1999 AEA Annual Meeting
Orlando, Florida

TM030314

(This paper is an adaptation of: **Assessing Student Outcomes**
Paper Presented at the Fourth Annual NISE Forum and Revised in September, 1999)

PRDE



RCSE COMM

Charting the Future of Assessment in Systemic Educational Reform: Teacher and School Principal Involvement in Evaluation and Assessment Use

Norma Dávila
University of Puerto Rico

Paper Presented at the 1999 AEA Annual Meeting
Orlando, Florida

(This paper is an adaptation of: **Assessing Student Outcomes**
Paper Presented at the Fourth Annual NISE Forum and Revised in September, 1999)

Overview

Student academic achievement is often the main area of interest for educators and policy makers within any discussion of systemic educational reform. These discussions are usually centered on traditional test scores that may or may not reflect what is important for reformers and educators yet, for many, they are the only available mechanism to demonstrate the impact of an initiative. Finding and designing alternative ways to measure student academic achievement within the new parameters of systemic educational reforms has been a major challenge for both evaluators and reformers who have searched together for answers to accountability questions. This paper presents the evolution of and the lessons learned from a research approach to assessment of student outcomes, specifically of student academic achievement, being used by the Puerto Rico Statewide Systemic Initiative (PR-SSI) which is one of the statewide systemic initiatives for Science and Mathematics sponsored by the National Science Foundation.

Definition of Outcomes and Outcome Variables

Weiss (1998) describes outcomes as "the end results of the program for the people it was intended to serve" (p.8) and further elaborates that outcomes are interchangeable with results and effects. Outcomes are certainly an end result of systemic educational reforms as well as of many other types of programs, but the nature and context of these initiatives requires a wider definition. For example, in systemic educational reforms, outcomes can be evident at the level of the classroom, school, district, or state. Evaluators of systemic educational reforms are usually interested in connections between different interventions and outcomes as well as in the factors that contributed to the occurrence of those outcomes.

Because of the additional dimensions of systemic educational reforms that differentiate these programs from other educational interventions, distinctions between outcome variables and outcomes need to be established. In systemic educational reform, an outcome variable is a quantity, dimension, or quality of the system subject to change because of the initiative. A systemic variable is an outcome variable that can be measured across the system such as student academic achievement in Science and Mathematics. In turn, an outcome for a systemic initiative is a change in an outcome variable directly attributable or likely attributable to the initiative such as improvements in student learning as a result of participation in standards-based instruction in Science and Mathematics.

Importance of Student Achievement Outcomes within Systemic Educational Reforms

The central focus of most systemic educational reforms the achievement of challenging academic standards that can be demonstrated through improvements in student academic achievement. Student academic achievement is interrelated with aspects of

the initiatives such as their visions of quality education, expectations of performance for participants, definitions of equity, and designs of professional development interventions among others. Further, student academic achievement is a concrete indicator of progress that is associated with other areas of student success such as college and job placement. Thus, systemic educational reforms are often expected to provide evidence of having an impact on student academic achievement as an indicator of the value added by the reforms. Consequently, evaluators face the challenge of choosing an appropriate data collection and reporting design that meets the needs of the initiatives and of their multiple stakeholders.

The Evolution of a Research Approach in the Assessment of Student Outcomes

Just like many other systemic educational reforms in Science and Mathematics, the PR-SSI's central focus is the student as an active learner (Shields, March, & Adelman, 1998). The PR-SSI fosters the holistic development of the students in preparation for their participation in the next century as illustrated in the constructivist principles that guide this reform; the PR-SSI envisions the teaching and learning process as bi-directional and interactive with the guidance of teachers within the context of school environments (Dávila, Vega & Rodríguez, 1996). A participatory-research approach was selected for the evaluation and assessment of the PR-SSI in general and for the assessment of student academic achievement in particular because: (1) the philosophy that guides this initiative emphasizes participant empowerment and the development of self-sustaining communities of learners; (2) the size and scope of the initiative require the involvement of increasing numbers of individuals; (3) the PR-SSI's reformers and participants possess expertise in a diversity of areas that can significantly contribute to the successful implementation of such a model; and (4) the literature available at the beginning of the initiative's implementation (i.e., 1991) clearly demonstrated a need for results of systemic educational reform based on research (Dávila, 1996).

Triangulation of results has been a major element of this design from the beginning of this reform. By comparing findings obtained using multiple quantitative and qualitative data collection strategies as suggested in the literature (Laguarda, Goldstein, Adelman & Zucker, 1998), the PR-SSI has identified trends and made pertinent mid-course corrections within its encompassing systemic strategy. The PR-SSI's participatory research evaluation and assessment design involved all the different areas being addressed by this comprehensive Science and Mathematics reform (see Dávila & Gómez, 1994; 1995; Dávila, Gómez & Vega, 1996 among others for specific examples). However, documenting and measuring student academic achievement was a major area of emphasis of this design because of (1) its importance within the larger context of systemic initiatives, and most importantly, (2) its value for the PR-SSI for decision-making purposes.

First Version of the PR-SSI's Model to Assess Student Academic Achievement

The first version of the model consisted of collecting and interpreting data at three different levels: (1) the classroom; (2) the initiative; and (3) the system (see Figure 1) (Puerto Rico Statewide Systemic Initiative, 1997). The description of each one of these levels follows.

As part of their professional development, Science and Mathematics teachers learn to use authentic assessment strategies such as open-ended questions, performance tasks, portfolios, and multiple choice questions that require higher order thinking skills to obtain information about student progress. Teachers use the results provided by these innovative strategies in their classrooms to (1) provide feedback to students about their performance and (2) modify their teaching, learning, and assessment practices. Teachers also translate these results into letter grades; schools provide grade distributions in terms of satisfactory (i.e., A's, B's, C's) and unsatisfactory (i.e., D's and F's) before and after their participation in the PR-SSI to identify trends in student academic achievement.

The initiative's staff developed a series of standards-based pre/post tests in Science and Mathematics to measure the value added by the systemic educational reform as part of the second level of the model. These tests included multiple-choice items that measure higher order thinking skills, open-ended questions, and performance tasks. Thus, assessment of student academic achievement was aligned at the classroom and initiative levels. Initially, all participating students took these assessments and, later, as the number of students and schools increased, representative samples of students were selected to represent their schools in the assessments.

The third level of the model consisted of external indicators of student progress for the overall K-12 system. The results of these tests provided other measures for the PR-SSI to "take the pulse" of the reform even though they were not fully aligned with the standards-based reform. These indicators included an adaptation and translation of the National Assessment of Educational Progress (NAEP) that was administered in 1994 in both Science and Mathematics to samples of participating PR-SSI students (i.e., lower socio-economic levels), students from private schools (i.e., middle and upper middle socio-economic levels), and students from non-participating public schools (i.e., lower socio-economic levels). They also included other tests designed by testing corporations and administered by the Puerto Rico Department of Education such as the SENDA and the Puerto Rican Competencies Test.

The first version of the model provided very useful information to the PR-SSI. However, as the needs of the initiative evolved, new ways to (1) look at student academic achievement; (2) provide specific formative feedback of student academic achievement to multiple players and stakeholders; and (3) design more mechanisms to drive the improvement of student learning in Science and Mathematics were imperative.

Second Version of the PR-SSI's Model to Assess Student Academic Achievement

The centerpiece of the second version of the model is the Science and Mathematics pre/post tests designed by the PR-SSI's staff in an alliance with The College Entrance Examination Board (CEEB) who provided technical expertise for their administration and analysis (see Figure 2) (Puerto Rico Statewide Systemic Initiative, 1998). The new tests were designed to measure achievement gains over the course of one year using public-released multiple-choice and open-ended items from NAEP and TIMSS. The tests were administered at the fourth, eighth, and eleventh grades; students from the 377 PR-SSI schools participated in this new assessment.

The new standards-based tests are scored using a scale equated with the TIMSS scale for item difficulty and student ability; a score of 500 in either scale equals the international average. By using an equivalent scale to that of TIMSS, student scores can be compared against national and international benchmarks of student performance that allow the PR-SSI to place the progress of its students within the larger global context (see Figure 3). Further, the results of these tests serve to guide revisions of the standards-based curricula being implemented in the classrooms that can be made by the teachers as well as by the Central PR-SSI staff.

Using Student Assessment Outcomes to Guide Teacher Professional Development

Another key element of the second version of the model is the teachers' participation in parallel assessments; their main purpose is to identify teachers' weaknesses in content that can be corrected through professional development. In this model, teachers receive sets of items not included in the tests administered to their students (but similar in approach and content) during a professional development session and are asked to respond to them anonymously. An item by item analysis of the distribution of their responses leads to a discussion of common misconceptions held by the teachers and of ways to correct these misconceptions. Lead teachers who provide direct academic and technical assistance to their peers participated in similar experiences led by university faculty before they designed and conducted these experiences for the teachers. The university faculty modelled the way to implement this model with the teachers and school principals; they designed the professional development experience for the participating teachers working in collaboration with the lead teachers.

The information provided by these analyses is another mechanism to refocus the initiative's professional development activities to address specific content needs of the teachers and of the lead teachers after they become aware of their needs in a non-threatening setting. The experience of the Central PR-SSI staff during these discussion sessions has been that the teachers often point out similarities between the content areas where they need additional professional development and the content areas where their students need additional instruction. Thus, by addressing the needs

of the teachers using these parallel assessments the PR-SSI is contributing to strengthen the quality of instruction at the classroom level.

Using External Criteria to Assess Student Outcomes

An external criterion now included in the PR-SSI's assessment of student academic achievement is the results of the college admissions tests administered by the CEEB. Since equating studies between the SAT and the CEEB Mathematics tests show a correlation 0.87, the PR-SSI can confidently compare the results of students in the Mathematics test of the CEEB with those of mainland students in the Mathematics test of the SAT (see Figure 3).

Another external criterion is the college admissions ratios to the University of Puerto Rico System which is the most competitive university system of the Island. College admissions ratios of PR-SSI participants are being analyzed by length of initiative intervention (i.e., intermediate school only vs. intermediate and high school). Distributions of chosen field of studies upon admissions are being analyzed in a similar way.

The evolution of the first and second versions of the PR-SSI's model to assess student academic achievement as an outcome of systemic educational reform show that considerable organizational learning has taken place within the PR-SSI. The following section addresses some of the lessons that the leadership of this reform has learned in the process of designing these models.

Sharing Results: Using Student Assessment Outcomes at the School Level

As soon as the results of the pre-tests are available, PR-SSI staff meet with school principals to discuss them and to guide their interpretation. The PR-SSI reports scores at the school level by item category and positions each school relative to the others within its geographical region. Each school principal receives the scores of his/her school and its relative position to the others, but does not receive any specific information about the other schools. The PR-SSI encourages school principal to take a critical look at the performance of their schools to identify areas in need of improvement as suggested by the item categories such as those proposed by the NCTM standards. The results of the post-tests are shared and discussed in a similar way.

By sharing the schools' results of the pre-tests by content area with school principals and teachers, the school can assume responsibility to improve student learning that can be demonstrated in the post-tests. For instance, teachers can refocus the content areas which they will reinforce in the classroom based on the needs of their students and can choose the most appropriate teaching and assessment strategies to meet those needs. As a result of these discussions, school principals become more aware of the academic needs of their students and, together with the teachers, can find ways to improve the teaching/learning environment of the schools to meet those needs. Thus, discussing

and analyzing these results is another way for the PR-SSI to have a direct impact on curricula and instruction as well as on the transformation of the school's teaching and learning culture and on the formation of communities of learners.

Lessons Learned

The process of designing the two versions of the assessment model required intense reflection and thinking by the leadership of the PR-SSI at multiple levels. Since the first version of the model had provided the initiative with very useful information over the years, it was difficult at first to make the decision to find another way to measure student academic achievement. However, the national exposure and dissemination of the TIMSS reports since 1997 was certainly a factor that prompted us to look for other alternatives more in tune with the evolving needs of the reform. Using public-released items from NAEP and TIMSS represented a major cost-saving step since the items had already been developed, but, without the vision and expertise of The College Entrance Examination Board , we would not have achieved the same results. At the same time, the PR-SSI staff is influencing the test design vision of this major player in education by emphasizing and modelling the use of national standards to guide test design. Further, the involvement and engagement of lead teachers, teachers, and school principals in the professional development exercises described above gave us pleasant surprises since they sincerely enjoy the experience of looking at their own performance and, most importantly, they grow professionally and personally in the process.

This model challenges traditional notions of participant involvement in evaluation because it requires: (1) direct contributions from participants to use results to improve teaching and learning; and (2) teamwork of evaluators and content specialists to lead the self-assessment process and its consequences. The model represents another step forward in the direction of evaluation and assessment ownership promoted by participatory evaluation designs.

Final Comments and Next Steps

One of the major challenges currently faced by evaluators and reformers who work with systemic educational reforms is the need for common metrics of student academic achievement. This is a recurrent theme in meetings sponsored by the National Science Foundation and it is evidently a high priority in the national educational reform agenda. We believe that the models presented in this paper can contribute to advance the design of such metrics.

For further information, please contact:

Norma Dávila, Ph.D.
Co-Principal Investigator, Puerto Rico Statewide Systemic Initiative
University of Puerto Rico
n_davila@upr1.upr.clu.edu

References

Dávila, N. (1996). Design and use of an interactive evaluation model to assess systemic change in Science and Mathematics education. Paper presented at the ASCD 51st. Annual Conference, New Orleans, Louisiana.

Dávila, N. & Gómez, M. (1994). Assessment of the impact of a new curriculum on systemic change. Paper presented at the AERA Annual Meeting, New Orleans, Louisiana. ERIC DOCUMENT 390 307

Dávila, N. & Gómez, M. (1995). Evaluation of school-based regional dissemination centers as scale-up mechanisms for systemic educational reform in Science and Mathematics. Paper presented at the AERA Annual Meeting, San Francisco, California. ERIC DOCUMENT 310 702

Dávila, N., Gómez, M. & Vega, I.Y. (1996). Evaluating the transformation of the teaching/learning culture of schools involved in systemic Science and Mathematics educational reform. Paper presented at the AERA Annual Meeting, New York, New York. ERIC DOCUMENT 395 803

Dávila, N., Vega, I.Y., & Rodríguez, J. (1996). Fomentando el desarrollo integral de los y las estudiantes mediante un currículo constructivista-interaccionista de Ciencias y Matemáticas: Implantación y resultados. Paper presented at the Eighth National Education and Thought Encounter, Ponce, Puerto Rico.

Laguarda, K.G., Goldstein, D.S., Adelman, N.E., & Zucker, A.A. (1998). Evaluation of the National Science Foundation's Statewide Systemic Initiative (SSI) Program. Assessing the SSI's impact on student achievement: An imperfect science. Menlo Park, CA: SRI International.

Puerto Rico Statewide Systemic Initiative (1997). Phase II Proposal.

Puerto Rico Statewide Systemic Initiative (1998). Performance Effectiveness Review, Report, December 1998.

Shields, P.M., Marsh, J.A., Adelman, N.E. (1998). Evaluation of NSF's Statewide Systemic Initiatives (SSI) Program: The SSI's impact on classroom practice. Menlo Park, CA: SRI International.

Weiss, C.H. (1998). Evaluation: Methods for studying programs and policies. Upper Saddle River, New Jersey: Prentice Hall.

Figure 1

Systemic Assessment of Student Performance Value Added by the Reform Three Tier Approach

Puerto Rican Competency Test in Science and Mathematics

Nationally recognized assessment experts assist in the development of a criterion-referenced, standard-based science assessment

Puerto Rico Assessment of Educational Progress (Adaptation of NAEP)

External Evaluation of Systemic Impact

Comparison of student progress in P.R. with respect to external criteria

Calibration of Program Assessment with National Assessment

Relative progress of SSI schools with respect to other schools in P.R.

Systemic Assessment of SSI in achieving Standards and SS&C / NCTM / NRC precepts

Program Level

Pretest - Postest to measure progress in achieving SSI educational goals of curriculum

Calibrate student progress against national and international standards with the assistance of national assessment experts using public released NAEP and TIMSS items

Value added by SSI Curriculum and school intervention

Calibration of Quality of Classroom Assessment with Program Assessment

Classroom Assessment

1. Portfolios and reflexives diaries
2. Open ended questions
3. Smart bubbles
4. Performance based assessment

Individual Student Performance Level

Figure 2

PR - SSI

Driving the PR-SSI Reform Process Through Standards-Based Testing and Benchmarking Against International Performance Indicators

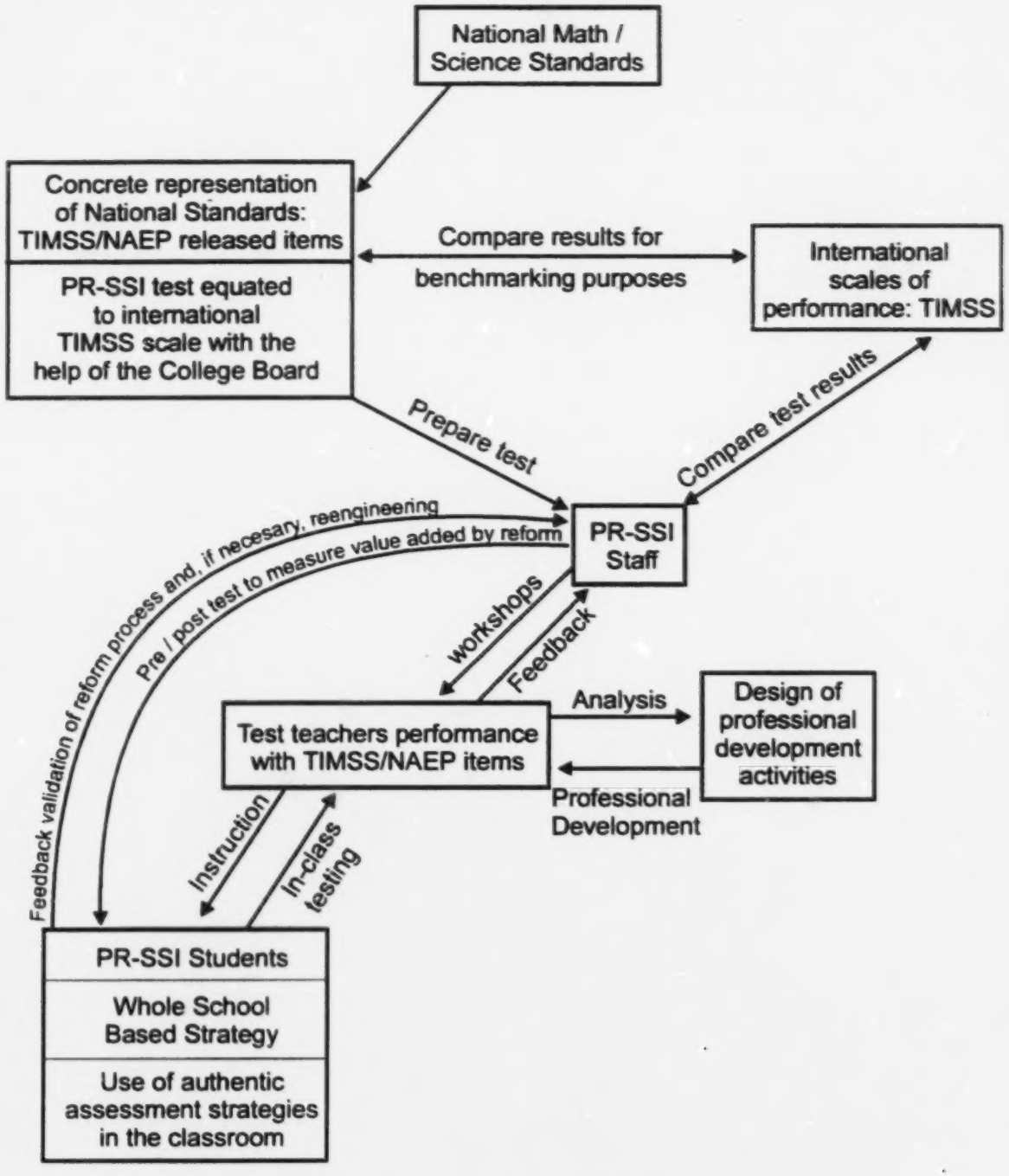
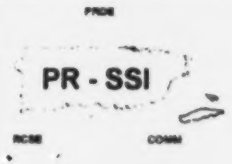

National Math / Science Standards
Concrete representation of National Standards: TIMSS/NAEP released items
PR-SSI test equated to international TIMSS scale with the help of the College Board
Compare results for benchmarking purposes
International scales of performance: TIMSS
Prepare test
Compare test results
PR-SSI Staff
Feedback validation of reform process and, if necesary, reengineering
Pre / post test to measure value added by reform
workshops
Feedback
Test teachers performance with TIMSS/NAEP items
Analysis
Design of professional development activities
Professional Development
Instruction
In-class testing
PR-SSI Students
Whole School Based Strategy
Use of authentic assessment strategies in the classroom

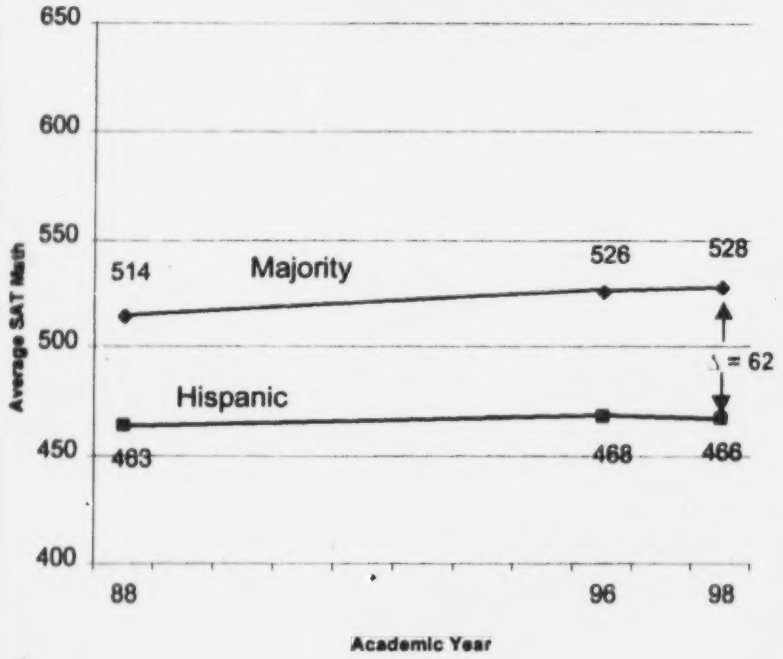
Figure 3



Longitudinal Study of Performance of Former PR-SSI Participants: Scores in College Admissions Tests (CEEB test and SAT test)

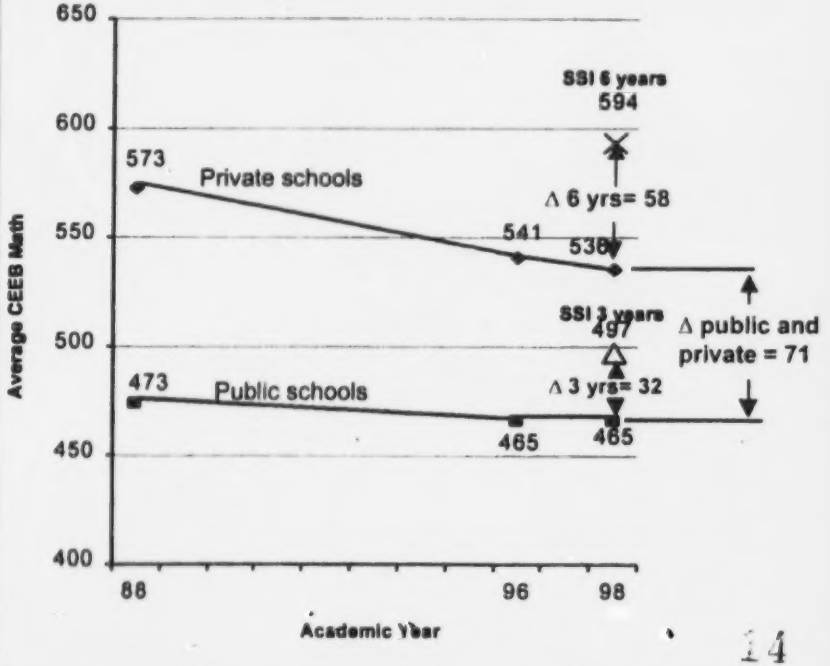
[Evidence of differences between continuous and interrupted reform interventions on student performance]

SAT Mathematics
Hispanic compared to Majority population
(Mainland)



13

College Board - Mathematics
(CEEB equating studies between CEEB and SAT tests show a correlation of 0.87 between both tests)



14

Private schools: students from families with middle and upper middle income backgrounds

Public schools: students from families with lower income backgrounds